Scheduling & Resource Management

www.gridforum.org
sched-wg@gridforum.org
www.nas.nasa.gov/~nitzberg/sched-wg/

Bill Nitzberg & Jennifer Schopf, co-chairs

Grid Forum Working Group Notes

Oct 19-21, 1999

Chicago, IL

Charter & Year 1 Goals

Initial Draft Presentation
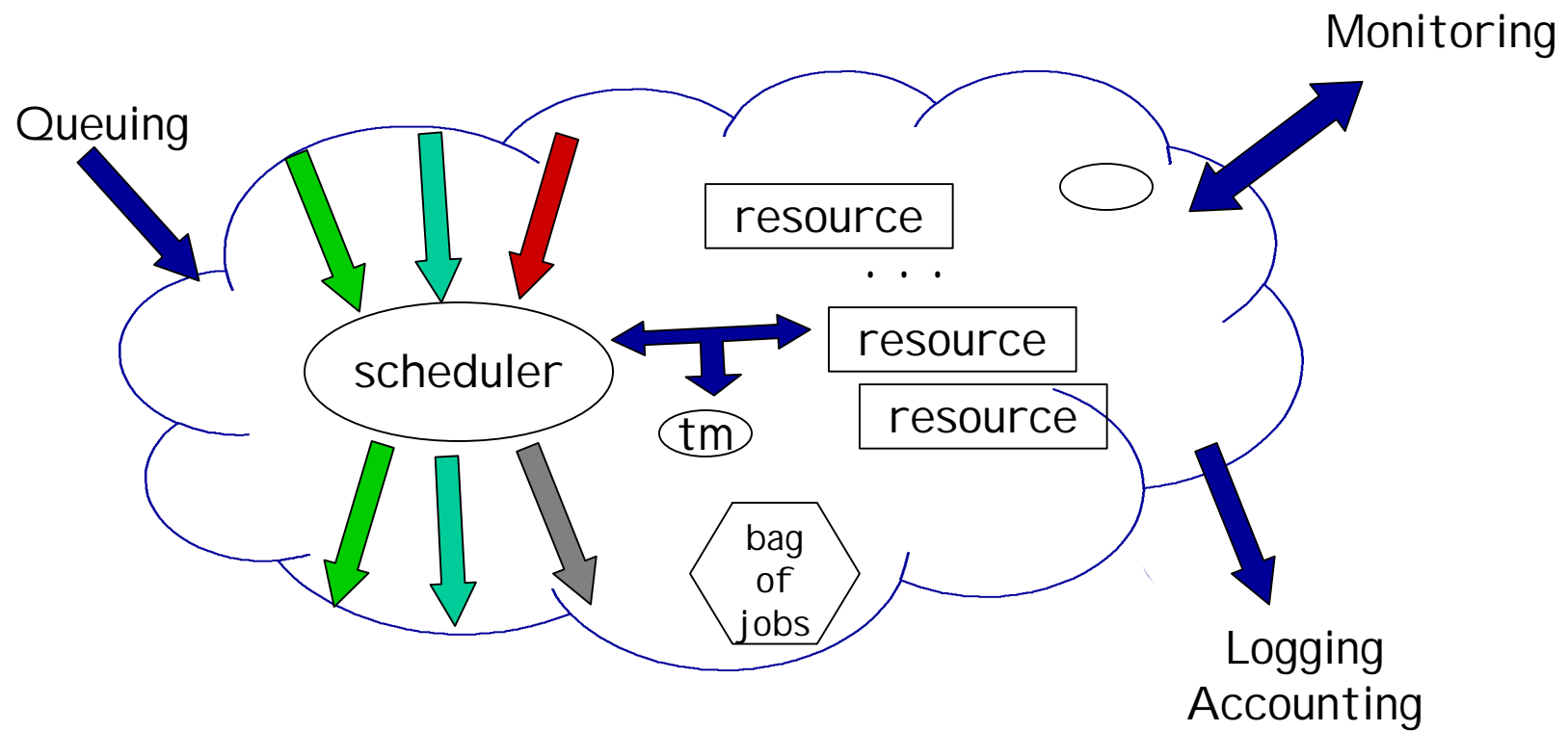Oct 19, 1999

Sched WG Charter (Proposed)

"Solve Grid Resource Management"

- Possible refinements:
  - focus on interfaces to promote interoperability
  - both policies and mechanisms
  - including: reservations, mobile code/spec., cost/performance/utilization trade-offs
  - standard interfaces between scheduler and…
    - meta/super-scheduler, peer scheduler, resource mgr, info. services, task mgr, monitoring, application
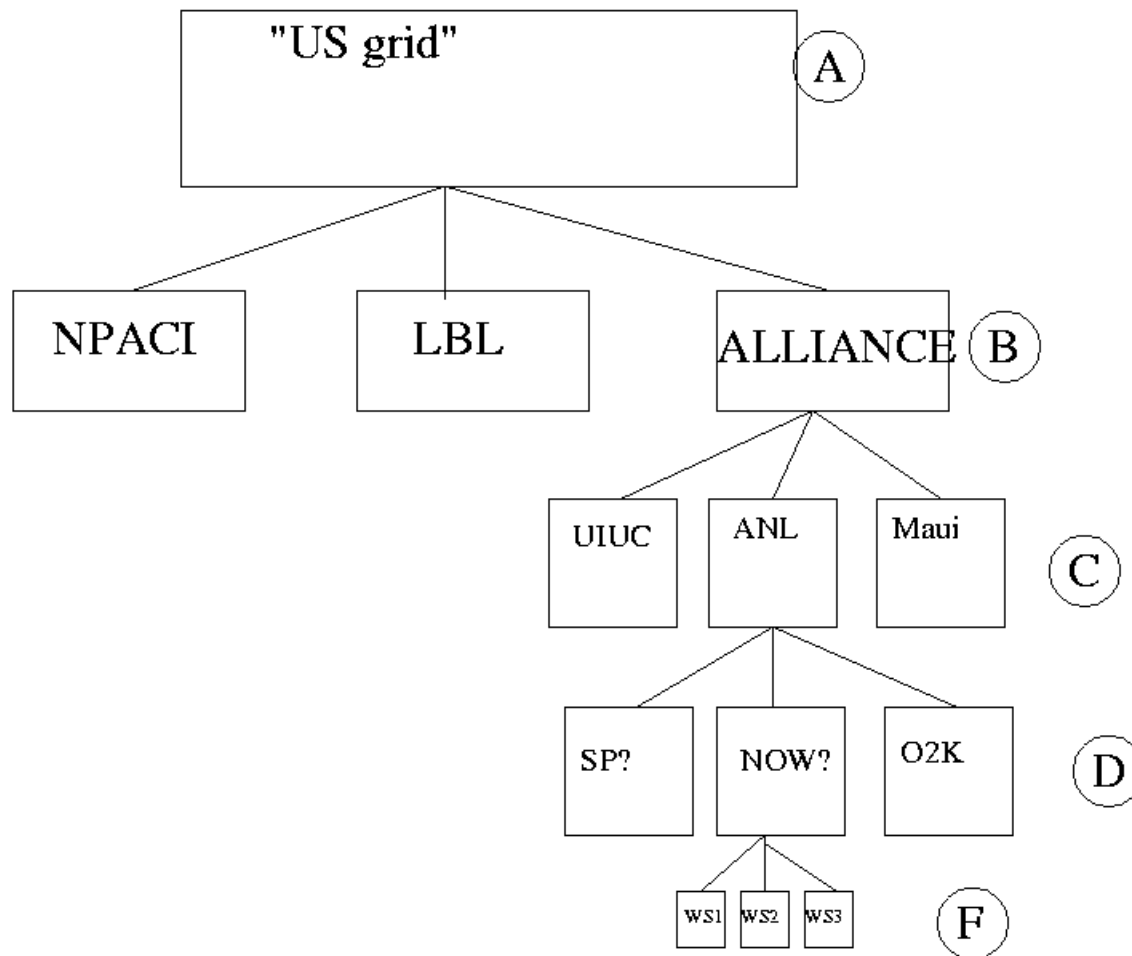
Sched WG 1 Year Goals (Proposed)

- Architecture
- Resource specification
- Advance reservations

- Others?
  - low-level job control API
  - workload trace format (RFC)
- Identify deferred items...

Architecture

Monitoring

Queuing

resource

. . .

resource

scheduler

tm

resource

bag
of
jobs

Logging
Accounting

Architecture

Grid Forum Sched WG Notes, Oct 1999, p. 6

Architecture

"submit job"

"job spec."

"scheduler"
POLICY

bag-o-jobs
bag-o-resources
POLICY

"queue
manager"

"job
manager"

. . .

"job
manager"

"queue
manager"

job
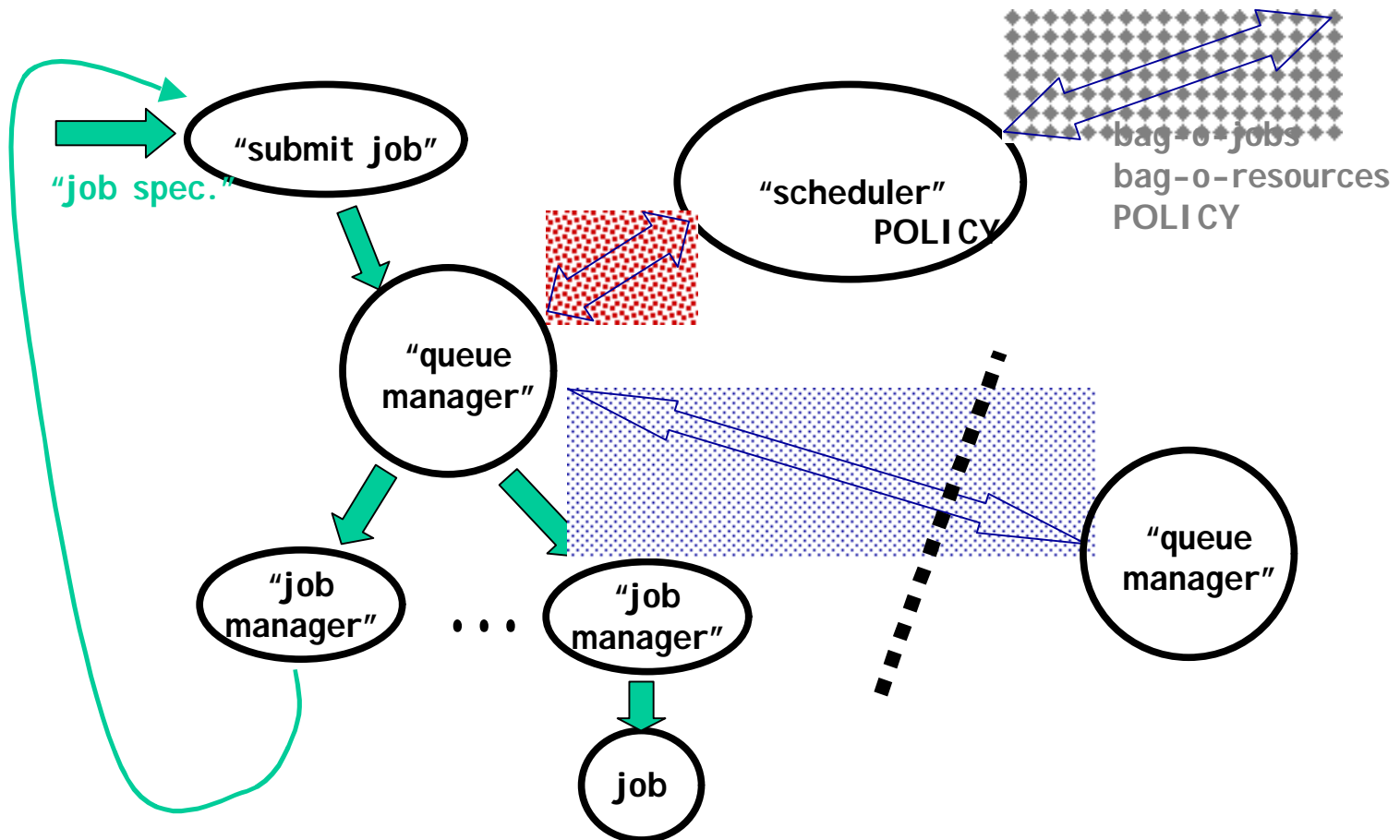
Resource Specification

- Terms
  - gathering existing terms from PBS, LSF, Globus, Legion, Condor, LL, Maui, ...

Advance Reservations

- Basic capabilities
  - (start, duration, end) wall-clock times
  - separate reservations from jobs
  - cancel / revoke
  - query / publish
- Prototypes
- What should go into the Info. Service?

Year 1 Goals Discussion

Oct 19, 1999

Year 1 Goal: Model / Architecture

- Agree on a model/architecture (e.g., picture with boxes & arrows); draft by GF3 (Feb '00)
  - terminology
  - state transitions that a job goes through (submission to execution)
  - list of resource specification names
  - sets of abstractions we're going to talk about when we're talking about scheduling on grids
  - examples of what an allocatable resource looks like

Grid Forum Sched WG Notes, Oct 1999, p. 11

Year 1 Goal: Advance Reservations

- Document desired advance reservations capabilities (i.e., from the advance reservation workshop); a draft white paper by GF3 (Feb '00)

    - API (C or command line) documented
    - encourage prototypes - want advance reservations across 3 different schedulers

Potential Year 1 Goals, cont.

- Standard for what a local allocatable resource looks like
  - What requirements do we need to have in order to allocate a local resource
  - What attributes should a local resource export
- Want ability to ask "how many resources are available right now and for how long?"
  - List of desired capabilities for "advanced" scheduling

Potential Year 1 Goals, cont.

- Workload trace format (RFC)
- "Standard" interface (API?) between
    - application & info. Source
    - application executable & resource manager
    - could be used to test the model
    - provides information categories

Sched WG
Breakout Meeting Notes

Oct 20, 1999

Plans for Today

- Review existing architecture pictures; discuss "common terms" & states

- Identify volunteers to post resource names/attributes from Globus, Legion, Condor, Load Leveler, Maui, ?

- Discuss process by which charter/goals will be refined

- Review advance reservation capabilities from May99 workshop as a starting point for the Grid Forum effort

Architecture Picture Needs...

- Resource discovery
  - includes both application and resource information

- Static placement vs. dynamic placement

- Feedback from the job manager to the scheduler

- Capture scheduling generic "things"
  - e.g., file staging, communication, ...

- Quality of service & accounting

- Admission control

Architecture

- Request groups with ready-made architectures of their systems post them to the list (along with an explanation of the boxes and arrows)

What is a "Job"?

- An entity that consumes resources to get things done
- A collection of "tasks"; a "task" is a collection of "task"s or a "pc + resources"
- POSIX def: a "session"... an "inescapable session"
- jobs are "scheduled"
- What characteristics of a "job" do we really need to define?

What is a "Job"?

- A "job" is a simple task, set of tasks, set of constraints, dependencies, resource annotations (e.g., bandwidth)

- A "simple job" is an atomic task with no constraints and no dependencies

- A job is something a scheduler schedules

What's a "scheduler"?

- A scheduler <u>maps</u> resources to jobs consistent with a set of constraints (e.g., time, policy)

Resource Names/Attributes
(the following will post to the list)

- Legion - Andrew Grimshaw
- Unicore - Dietmar Erwin
- Globus - Gregor von Laszewski
- Maui - Quinn Snell
- GRD/Codine - Craig Stair
- NQE - Sam Watters
- Unassigned: Load Leveler, DPSS, ...
- Already posted to list: PBS, LSF/NCSA

Summary of Actions

- Draft model/architecture "RFC" (inc. attributes) for GF3 (Feb '00)
  - collect & merge resource names/attributes from existing systems
  - collect & merge existing architecture pictures
- Draft advance reservation capabilities + model "RFC" for GF3 (Feb '00)
- Refine charter via email
  - process is to suggest actual textual changes to existing charter on the list

Grid Forum Sched WG Notes, Oct 1999, p. 23